STOCK PRICE PREDICTION

PHASE 2 - INNOVATION

INTRODUCTION:

Stock price prediction refers to the application of advanced technologies and novel approaches to forecast future stock prices or market trends. It involves leveraging innovative methods, data sources, and computational techniques to improve the accuracy and efficiency of predicting how the prices of financial assets, such as stocks and securities, will behave in the future

STEP 1: DATA COLLECTION AND PREPROCESSING

DATA COLLECTION:

Data collection is the process of systematically gathering and acquiring relevant information or data from various sources to support analysis, modeling, and insights generation.

PREPROCESSING:

Preprocessing is the set of procedures and techniques used to clean, transform, and prepare raw data for analysis, modeling, or machine learning, ensuring it is in a suitable format for further processing.

STEP 2: FEATURE ENGINEERING

CREATING INFORMATIVE FEATURES:

It is the process of generating new features from the existing dataset or domain knowledge that can enhance the predictive power of machine learning models. This can include mathematical transformations, aggregations, interactions between variables, and the creation of new variables that capture relevant information.

DIMENSIONALITY REDUCTION:

Feature engineering also includes techniques to reduce the dimensionality of data while preserving essential information. Methods like Principal Component Analysis (PCA) or feature selection algorithms help eliminate irrelevant or redundant features, making models more efficient and interpretable.

STEP 3: MODEL SELECTION

- Train and evaluate each candidate regression model using appropriate evaluation metrics. Using techniques like cross-validation to assess the models' performance on validation data.

- Consider how well each model handles the specific challenges of stock price prediction, such as capturing volatility and seasonality.

STEP 4: MODEL TRAINING

- Train the selected regression model using the training data. The model learns to make predictions by finding the relationships between the input features and the target variable.

- The training process involves adjusting the model's internal parameters iteratively to minimize the chosen loss function (e.g., Mean Squared Error) on the training data.

STEP 5: MODEL EVALUATION

EVALUATION METRICS:

Select appropriate evaluation metrics to measure the accuracy and reliability of your regression model's prediction.

MEAN ABSOLUTE ERROR (MAE):

Measures the average absolute difference between the predicted and actual values. Smaller MAE values indicate better accuracy

MEAN SQUARED ERROR (MSE):

Measures the average squared difference between the predicted and actual values. It penalizes larger errors more heavily and provides insight into the model's ability to capture variations in the data.

ROOT MEAN ERROR (RMSE):

The square root of MSE, which provides an interpretable metric in the same units as the target variable.

R-SQUARED (R^2):

Also known as the coefficient of determination, R^2 quantifies the proportion of variance in the target variable explained by the model. A higher R^2 indicates a better fit to the data.

STEP 6: DEPLOYMENT AND PREDICTION

- Deploy the chosen regression model to predict the stock price prediction.
- Assess the accuracy and reliability of the model's predictions on the new data.

CONCLUSION:

In phase 2 conclusion, we will summarize the key findings and insights from the advanced regression technique. We will reiterate the impact of these techniques on improving the accuracy and robustness of stock price prediction.